ITERATIVE NONLINEAR LEAST SQUARES ALGORITHMS FOR DIRECT RECONSTRUCTION OF PARAMETRIC IMAGES FROM DYNAMIC PET

Guobao Wang and Jinyi Qi

Department of Biomedical Engineering, University of California, Davis, CA 95616

ABSTRACT

Indirect and direct methods have been developed for reconstructing parametric images from dynamic PET data. Indirect methods are relatively simple and easy to implement because the reconstruction and kinetic modeling are performed in two separate steps. Direct methods estimate the parametric images directly from the dynamic PET data and are statistically more efficient, but the algorithms are often difficult to implement. This paper presents a simple, monotonically convergent iterative algorithm for direct reconstruction of parametric images. Each iteration of the proposed algorithm consists of two separate steps: reconstruction of dynamic images followed by a pixel-wise weighted nonlinear least squares fitting. This algorithm resembles the empirical iterative implementation of the indirect approach, but converges to the solution of the direct formulation.

Index Terms— Image reconstruction, tracer kinetic modeling, parametric imaging, dynamic PET

1. INTRODUCTION

Parametric imaging using dynamic positron emission tomography (PET) provides important information for biological research and clinical diagnosis. To obtain a parametric image, a typical approach is to reconstruct a sequence of emission images from the measured projection data first, and then to fit the time activity curve (TAC) at each pixel to a linear or nonlinear kinetic model. To obtain an efficient estimate, the noise distribution of the reconstructed emission images should be modeled in the kinetic analysis. However, exact modeling of the noise distribution in emission images reconstructed by iterative methods is difficult because the noise is space-variant and object-dependent. Usually the space-variant noise variance and correlations between pixels are simply ignored in the kinetic analysis, which leads to sub-optimal results. Direct reconstruction of parametric images from the raw projection data solves this problem by combining kinetic modeling and emission image reconstruction into a single formula. It allows accurate modeling of the detector response of PET systems and Poisson noise statistics in data. It has been shown that direct reconstruction can achieve better bias-variance tradeoff compared with indirect methods [1, 2].

One drawback of direct reconstruction is that the optimization algorithms are usually complex when compared to indirect methods, because the kinetic parameters are involved in the reconstruction formulation nonlinearly [1]. For indirect methods, optimization algorithms for both image reconstruction and nonlinear least squares (NLS) kinetic fitting have been well developed. To take advantages of these results, here we develop a simple, easy to implement and montonically convergent algorithm for direct reconstruction. Each iteration of the proposed algorithm consists of two separate steps: reconstruction of dynamic images and a pixel-wise weighted NLS kinetic fitting. The algorithm resembles the empirical iterative implementation of the indirect approach (e.g. [3]), but pursuits the solution of direct reconstruction. Since NLS is used iteratively in the proposed algorithm, we call it *iterative NLS (INLS)*.

2. RECONSTRUCTION OF PARAMETRIC IMAGES

2.1. Indirect Methods

In indirect methods, the reconstruction of emission images and kinetic modeling are treated separately. Efficient algorithms for both image reconstruction and kinetic modeling have been developed.

2.1.1. Emission Image Reconstruction

PET data are modeled as a collection of independent Poisson random variables with the expectation $\bar{\boldsymbol{y}}_m$ in frame m related to the image $\boldsymbol{x}_m$ through an affine transform

$$\bar{\boldsymbol{y}}_m = \boldsymbol{P}\boldsymbol{x}_m + \boldsymbol{r}_m, \tag{1}$$

where $\boldsymbol{P} \in I\!R^{M\times N}$ is the detection probability matrix with element (i, j) being the probability of detecting an event originated in voxel j by detector pair i, and $\boldsymbol{r}_m \in I\!R^M$ is the expectation of scattered and random events in the mth frame. The log-likelihood function of the dynamic data set, omitting constants that are independent of $\boldsymbol{x}$, is

$$\mathcal{L}(\boldsymbol{y}|\boldsymbol{x}) = \sum_m \sum_i y_{im} \log \bar{y}_{im} - \bar{y}_{im}, \tag{2}$$

This work was supported by NIH under grant R01 EB00194.

where $\boldsymbol{y} = \{\boldsymbol{y}_m\}$ denotes the measured dynamic sinograms and $\boldsymbol{x} = \{\boldsymbol{x}_m\}$ denotes the unknown dynamic images.

The estimation of the dynamic emission images $\boldsymbol{x}$ can be achieved by the penalized maximum likelihood method that maximizes an objective function as

$$\hat{\boldsymbol{x}} = \arg\max_{\boldsymbol{x}\geq\boldsymbol{0}} \mathcal{L}(\boldsymbol{y}|\boldsymbol{x}) - \beta\mathcal{U}(\boldsymbol{x}), \tag{3}$$

where $\mathcal{U}(\boldsymbol{x})$ is a roughness penalty for regularizing the noise and β is the regularization parameter that controls the tradeoff between the resolution and noise. Commonly used penalty functions can be written in the following form

$$\mathcal{U}(\boldsymbol{x}) = \sum_m \alpha_m U(\boldsymbol{x}_m),\ U(\boldsymbol{x}_m) = \sum_{k=1}^{K} \psi_k\big([\boldsymbol{D}\boldsymbol{x}_m]_k\big), \tag{4}$$

where α_m is the weighting factor of frame m, usually determined based on the noise level in the frame. $\boldsymbol{D} \in \mathbb{R}^{K\times N}$ is a sparse neighborhood differentiation matrix whose kth row has nonzero elements corresponding to the pixels that form the kth clique, and $\psi_k(\cdot)$ is the potential function defined on the cliques within this neighborhood system.

2.1.2. Tracer Kinetic Modeling

Tracer kinetic behaviors in dynamic PET imaging are often described by compartmental models which mathematically can be represented by a set of ordinary differential equations,

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{c}(t) = \boldsymbol{K}\boldsymbol{c}(t) + \boldsymbol{L}\boldsymbol{u}(t) \tag{5}$$

where $\boldsymbol{c}(t)$ is a column vector representing the activity concentration of different tissue compartments at time t, $\boldsymbol{K}$ and $\boldsymbol{L}$ are the kinetic parameter matrices which comprise of various rate constants, and $\boldsymbol{u}(t)$ denotes the system input.

The quantity that PET measures is the total concentration

$$C_{\mathrm{T}}(t) = (1 - f_v)\mathbf{1}'\boldsymbol{c}(t) + f_v C_{wb}(t), \tag{6}$$

where $\mathbf{1}$ is the all-one vector, f_v is the fractional volume of the blood in the tissue, and $C_{wb}(t)$ is the tracer concentration in the whole blood. In practice, PET data are binned into discrete time frames. Thus the measured quantity in frame m is the average concentration

$$z_m(\boldsymbol{\kappa}) = \frac{1}{\Delta t_m}\int_{t_{m,s}}^{t_{m,e}} C_{\mathrm{T}}(\tau)e^{-\lambda\tau}\mathrm{d}\tau, \tag{7}$$

where $t_{n,s}$ and $t_{n,e}$ denote the starting and ending time of frame n, respectively, $\Delta t_m = t_{m,e} - t_{m,s}$, λ is the decay constant of the radiotracer, and $\boldsymbol{\kappa}$ contains the kinetic parameters to be estimated.

Given a measured TAC, $\hat{\boldsymbol{z}} = \{\hat{z}_m\}$, kinetic analysis is to estimate the rate constants in $\boldsymbol{K}$ and $\boldsymbol{L}$, and f_v, which is usually accomplished by using a NLS formulation

$$\hat{\boldsymbol{\kappa}} = \arg\min_{\boldsymbol{\kappa}} \sum_m w_m\Big(\hat{z}_m - z_m(\boldsymbol{\kappa})\Big)^2, \tag{8}$$

where $\{w_m\}$ are the weighting factors. A usual choice of w_m is Δt_m. A Levenberg-Marquart algorithm is commonly used to solve (8).

2.2. Direct Methods

Direct methods combine the kinetic model and image reconstruction into a single formula. Let $\boldsymbol{\kappa}_j$ denote the kinetic parameters for pixel j and $\boldsymbol{\kappa} = \{\boldsymbol{\kappa}_j\}$, the log-likelihood function is,

$$\mathcal{L}(\boldsymbol{y}|\boldsymbol{\kappa}) = \sum_{i,m} y_{im}\log\bar{y}_{im}(\boldsymbol{\kappa}) - \bar{y}_{im}(\boldsymbol{\kappa}). \tag{9}$$

The expectation of the data is now a nonlinear transform of $\boldsymbol{\kappa}$,

$$\bar{y}_{im}(\boldsymbol{\kappa}) = \sum_j p_{ij}x_m(\boldsymbol{\kappa}_j) + r_{im}, \tag{10}$$

where $x_m(\boldsymbol{\kappa}_j)$ denotes the image intensity at pixel j in the frame m due to the kinetics $\boldsymbol{\kappa}_j$. It relates to the time activity curve $z_m(\boldsymbol{\kappa}_j)$ via

$$x_m(\boldsymbol{\kappa}_j) = z_m(\boldsymbol{\kappa}_j)\Delta t_m. \tag{11}$$

The solution of the direct reconstruction is

$$\hat{\boldsymbol{\kappa}} = \arg\max_{\boldsymbol{\kappa}} \mathcal{L}(\boldsymbol{y}|\boldsymbol{\kappa}) - \beta\mathcal{U}\big(\boldsymbol{x}(\boldsymbol{\kappa})\big). \tag{12}$$

where the penaly term $\mathcal{U}\big(\boldsymbol{x}(\boldsymbol{\kappa})\big)$ has a similar form as (4). Here the regularization is applied on the image intensity.

3. THE PROPOSED ALGORITHM

The proposed algorithm resembles an iterative implementation of the indirect method. There are two steps at each iteration:
(1) Image reconstruction

$$\hat{x}_{jm}^{n+1} = x_m(\boldsymbol{\kappa}_j^n) + \frac{g_{jm}^n}{w_{jm}^n}, \tag{13}$$

(2) Kinetic fitting

$$\hat{\boldsymbol{\kappa}}_j^{n+1} = \arg\max_{\boldsymbol{\kappa}_j} -\sum_m w_{jm}^n\Big(x_m(\boldsymbol{\kappa}_j) - \hat{x}_{jm}^{n+1}\Big)^2, \tag{14}$$

where g_{jm}^n denotes the gradient of pixel j in frame m at the nth iteration, and w_{jm}^n is a weighting factor determined by the specific algorithm (explained below).

The advantage of the proposed algorithm is that any existing NLS algorithm for kinetic fitting can be directly applied in the second step with the selected weighting factors. It is worth noting that the weights in (14) are determined by the reconstruction algorithm, as compared to that the weights in the NLS fitting in indirect methods (including their empirical

iterative implementations) are selected by users. Another distinction is that it is not necessary to run the NLS fitting in (14) till convergence at each iteration. Here we use two iterations of the Levenberg-Marquart algorithm. Therefore, each iteration of the direct reconstruction takes about the same amount of time as that for one iteration of the dynamic image reconstruction and two iterations of the Levenberg-Marquart algorithm.

In the following, we derive the expressions of g_{jm}^n and w_{jm}^n in (13) using the optimization transfer principle [4].

3.1. Paraboloidal Surrogate Function

We use the paraboloidal surrogate function proposed by Fessler *et al* in [5] to approximate the Poisson log-likelihood function and a quadratic surrogate function to approximate the penalty function at each iteration. The overall surrogate function for the penalized log-likelihood in (12) is

$$\mathcal{Q}(\boldsymbol{\kappa}|\boldsymbol{\kappa}^n) = \sum_m (\boldsymbol{g}_m^n)' \Delta \boldsymbol{x}_m - \frac{1}{2} \Delta \boldsymbol{x}_m' \boldsymbol{W}_m^n \Delta \boldsymbol{x}_m, \quad (15)$$

where $\Delta \boldsymbol{x}_m = \boldsymbol{x}_m(\boldsymbol{\kappa}) - \boldsymbol{x}_m(\boldsymbol{\kappa}^n)$, $\boldsymbol{g}_m^n$ is the gradient vector of the penalized log-likelihood, and $\boldsymbol{W}_m^n$ is the curvature matrix. $\boldsymbol{g}_m^n$ and $\boldsymbol{W}_m^n$ are given by

$$\boldsymbol{g}_m^n = \boldsymbol{P}' \Big(\frac{\boldsymbol{y}_m}{\bar{\boldsymbol{y}}_m(\boldsymbol{\kappa}^n)} - \mathbf{1} \Big) - \beta_m \boldsymbol{\mathcal{D}}[\boldsymbol{\omega}_m^n] \boldsymbol{D} \boldsymbol{x}_m(\boldsymbol{\kappa}^n) \quad (16)$$

$$\boldsymbol{W}_m^n = \boldsymbol{P}' \boldsymbol{\mathcal{D}}[\boldsymbol{c}_m^n] \boldsymbol{P} + \beta_m \boldsymbol{D}' \boldsymbol{\mathcal{D}}[\boldsymbol{\omega}_m^n] \boldsymbol{D} \quad (17)$$

where $\beta_m = \beta \alpha_m$ and $\boldsymbol{\mathcal{D}}[\cdot]$ denotes a diagonal matrix. $\boldsymbol{c}_m^n$ is the curvature of the surrogate function for the log-likelihood function and $\boldsymbol{\omega}_m^n$ the curvature of the surrogate function for the penalty function. They both can be calculated analytically [5].

3.2. Coordinate Ascent

To estimate the iterate of $\boldsymbol{\kappa}$ by maximizing the surrogate function $\mathcal{Q}(\boldsymbol{\kappa}|\boldsymbol{\kappa}^n)$, a pixel-wise coordinate ascent (CA) algorithm can be applied

$$\hat{\boldsymbol{\kappa}}_j^{n+1} = \arg\max_{\boldsymbol{\kappa}_j} \mathcal{Q}_j(\boldsymbol{\kappa}_j|\boldsymbol{\kappa}^n, \hat{\boldsymbol{\kappa}}), \quad (18)$$

$$\mathcal{Q}_j(\boldsymbol{\kappa}_j|\boldsymbol{\kappa}^n, \hat{\boldsymbol{\kappa}}) = \sum_m \hat{g}_{jm}^n \Delta x_{jm} - \frac{1}{2} \hat{w}_{jm}^n \Delta x_{jm}^2, \quad (19)$$

where $\Delta x_{jm} = x_m(\boldsymbol{\kappa}_j) - x_m(\boldsymbol{\kappa}_j^n)$ and

$$\hat{g}_{jm}^n = g_{jm}^n + \sum_{l \neq j} [\boldsymbol{W}_m^n]_{jl} \big(x_m(\hat{\boldsymbol{\kappa}}_l) - x_m(\boldsymbol{\kappa}_l^n) \big), \quad (20)$$

$$\hat{w}_{jm}^n = [\boldsymbol{W}_m^n]_{jj} = \sum_i c_{im}^n p_{ij}^2 + \beta_m \sum_{l \in \mathcal{N}_j} \omega_{jlm}^n, \quad (21)$$

where $\mathcal{N}_j$ denotes the neighbors of the jth pixel and ω_{jlm}^n is the element of $\boldsymbol{\omega}_m^n$ that is associated with the clique formed by pixels j and l.

Maximizing $\mathcal{Q}_j(\boldsymbol{\kappa}_j|\boldsymbol{\kappa}^n, \hat{\boldsymbol{\kappa}})$ with respect to $\boldsymbol{\kappa}_j$ is equivalent to

$$\hat{\boldsymbol{\kappa}}_j^{n+1} = \arg\max_{\boldsymbol{\kappa}_j} -\frac{1}{2} \sum_m \hat{w}_{jm}^n \Big(x_m(\boldsymbol{\kappa}_j) - \hat{x}_{jm}^{n+1} \Big)^2, \quad (22)$$

where

$$\hat{x}_{jm}^{n+1} = x_m(\boldsymbol{\kappa}_j^n) + \frac{\hat{g}_{jm}^n}{\hat{w}_{jm}^n}. \quad (23)$$

3.3. Separable Paraboloid

The CA-INLS algorithm requires column-access of the system matrix $\boldsymbol{P}$, which is inefficient for factorized system matrices. By using the convex property of (15), we can get a separable surrogate $\mathcal{Q}^s(\boldsymbol{\kappa}|\boldsymbol{\kappa}^n)$

$$\mathcal{Q}^s(\boldsymbol{\kappa}|\boldsymbol{\kappa}^n) = \sum_m (\boldsymbol{g}_m^n)' \Delta \boldsymbol{x}_m - \frac{1}{2} \Delta \boldsymbol{x}_m' \boldsymbol{\mathcal{D}}[\boldsymbol{w}_m^n] \Delta \boldsymbol{x}_m, \quad (24)$$

where $\boldsymbol{w}_m^n$ is given by

$$w_{jm}^n = \sum_i c_{im}^n p_{ij} p_i + 2\beta_m \sum_{l \in \mathcal{N}_j} \omega_{jlm}^n \quad (25)$$

where $p_i = \sum_j p_{ij}$.

The new surrogate is composed of separable paraboloids (SP). Maximizing $\mathcal{Q}^s(\boldsymbol{\kappa}|\boldsymbol{\kappa}^n)$ with respect to $\boldsymbol{\kappa}$ is equivalent to

$$\hat{\boldsymbol{\kappa}}_j^{n+1} = \arg\max_{\boldsymbol{\kappa}_j} -\frac{1}{2} \sum_m w_{jm} \Big(x_m(\boldsymbol{\kappa}_j) - \hat{x}_{jm}^{n+1} \Big)^2, \quad (26)$$

where

$$\hat{x}_{jm}^{n+1} = x_m(\boldsymbol{\kappa}_j^n) + \frac{g_{jm}^n}{w_{jm}^n}. \quad (27)$$

Convergence: Both the CA-INLS and SP-INLS algorithms are monotonically convergent. However, because of the nonlinear relationship between the dynamic PET data and kinetic rates, the log-likelihood function $\mathcal{L}(\boldsymbol{y}|\boldsymbol{\kappa})$ is non-concave with respect to $\boldsymbol{\kappa}$. As a result, both algorithms can only guarantee convergence to a local optimum.

4. SIMULATION RESULTS

We used computer simulations to validate the proposed algorithms based on a brain phantom composed of gray matter, white matter and a small tumor inside the white matter (Fig. 1). Different FDG kinetics under a three-compartment model were simulated for different regions. The scanning sequence consists of 24 frames over a period of 90 minutes. The TACs were integrated for each frame and forward projected to generate dynamic sinograms. Poisson noise was then added, which resulted in a total number of events over the 90 minutes equal to 50M. Forty noisy datasets were generated and processed independently by the direct method and indirect

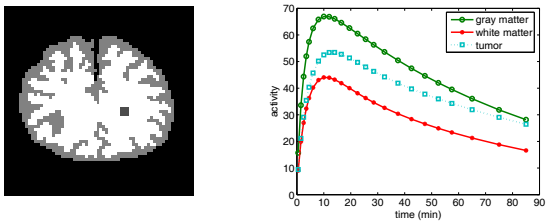
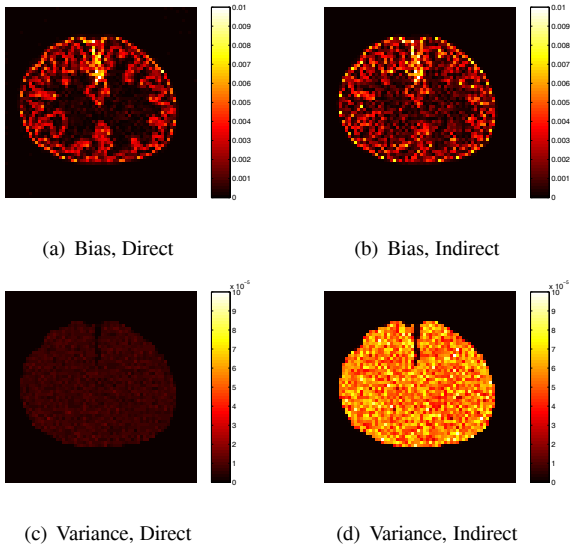


Fig. 1. The phantom and the noise-free regional time activity curves used in the simulation.



(a) Bias, Direct (b) Bias, Indirect

(c) Variance, Direct (d) Variance, Indirect

Fig. 2. The bias and variance images of the direct and indirect algorithms.

method to estimate the kinetic parameters. Here we focus on the estimation of the influx rate, a parameter that is related to the glucose metabolic rate.

Fig. 2 shows the bias and variance images of the influx rate images generated by the indirect and direct algorithms (CA-INLS) with quadratic penalty functions. The regularization parameters in the two approaches were adjusted to achieve similar bias (Fig. 2(a) and (b)). Obviously the direct algorithm results in less variance than the indirect algorithm (Fig. 2(c) and (d)).

Fig. 3 compares the tradeoff between bias and standard deviation of the influx rate estimated by the direct and indirect methods for different regions in the brain phantom. The bias-std curves were obtained by varying the regularization parameters. The results show that the proposed INLS method results in less bias at all the noise levels compared with the indirect method.

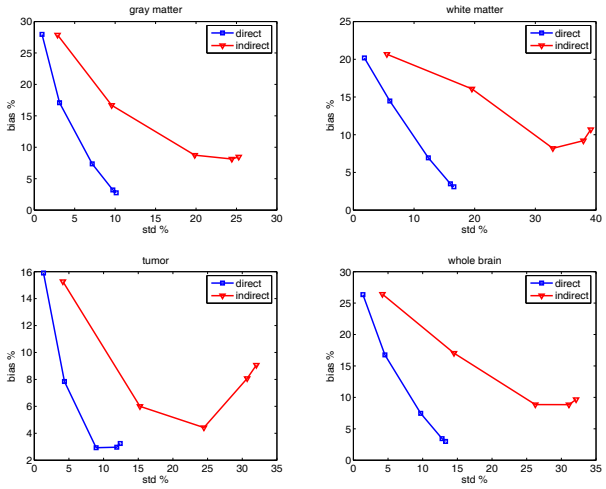


Fig. 3. The bias versus standard deviation tradeoff curves of the direct and indirect algorithms by varying the regularization parameters.

5. CONCLUSIONS

This paper develops a class of iterative NLS algorithms for direct reconstruction of parametric images. Each iteration of the proposed algorithm consists of one image update followed by a NLS fitting. The algorithms are simple, easy to implement and monotonically convergent.

6. REFERENCES

[1] M. E. Kamasak, C. A. Bouman, E. D. Morris, and K. Sauer, "Direct reconstruction of kinetic parameter images from dynamic PET data," *IEEE Trans. Med. Im.*, vol. 24, no. 5, pp. 636–650, 2005.

[2] G. B. Wang, L. Fu, and J. Qi, "Maximum a posteriori reconstruction of the patlak parametric image from sinograms in dynamic PET," *Phys. Med. Biol.*, vol. 53, no. 3, pp. 593–604, 2008.

[3] A. J. Reader, Matthews J. C., and Sureau F. C. *et al*, "Iterative kinetic parameter estimation within fully 4D PET image reconstruction," *IEEE Nucl. Sci. Symp. Med. Im. Conf.*, vol. 3, pp. 1752–1756, 2006.

[4] K. Lange, D. R. Hunter, and I. Yang, "Optimization transfer using surrogate objective functions," *J Computational and Graphical Statistics*, vol. 9, no. 1, pp. 1–20, 2000.

[5] J. A. Fessler and H. Erdogan, "A paraboloidal surrogates algorithm for convergent penalized-likelihood emission image reconstruction," *IEEE Nucl. Sci. Symp. Med. Im. Conf.*, vol. 2, pp. 1132–5, 1998.